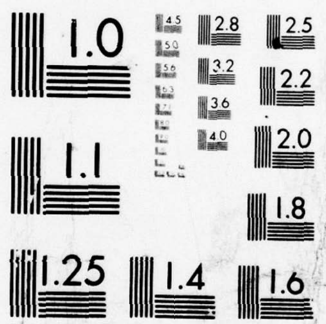
AD-A078 488 MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR INFORMA--ETC F/G 12/1
SOME PROPERTIES OF THE DUAL ADAPTIVE STOCHASTIC CONTROL ALGORIT--ETC(U)
AUG 79 P DERSIN , M ATHANS , D A KENDRICK AFOSR-77-3321
UNCLASSIFIED LIDS-P-936 AFOSR-TR-79-1254 NL

1 OF 1
AD
A078488

END
DATE
FILMED
1 — 80
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

LEVEL II

2

19 REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 18 AFOSR-TR- 79-1254	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): 6 SOME PROPERTIES OF THE DUAL ADAPTIVE STOCHASTIC CONTROL ALGORITHM	5. TYPE OF REPORT & PERIOD COVERED: 9 Interim rept.
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): 10 Pierre Dersin, Michael Athans, David A. Kendrick	8. CONTRACT OR GRANT NUMBER(s): 15 AFOSR-77-3321
9. PERFORMING ORGANIZATION NAME AND ADDRESS: MIT Laboratory for Information & Decision Systems Cambridge, MA 02139	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 16 2304/A1
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332	12. REPORT DATE: 11 August 1979 17 A1
	13. NUMBER OF PAGES: 30
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 12 31	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report) 14 LIDS-P-936

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The purpose of this paper is to compare analytically the properties of the suboptimal dual adaptive stochastic control algorithm when the plant dynamics contain multiplicative white noise parameters. A simple scalar example is used for this analysis.

410 950

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AFOSR-TR- 79-1254

August, 1979 LIDS-P-936

SOME PROPERTIES OF THE DUAL ADAPTIVE STOCHASTIC CONTROL ALGORITHM*

by

Pierre Dersin
Michael Athans
David A. Kendrick

Laboratory for Information and Decision Systems
Room 35-308
Massachusetts Institute of Technology
Cambridge, Mass. 02139

ABSTRACT

The purpose of this paper is to compare analytically the properties of the suboptimal dual adaptive stochastic control algorithm when the plant dynamics contain multiplicative white noise parameters. A simple scalar example is used for this analysis.

79 12 10 186

*Research supported by AFOSR Grant 77-3321[illegible].

Submitted to the IEEE Transactions on Automatic Control.

1. INTRODUCTION

Most stochastic optimal control problems are not amenable to a solution through the stochastic dynamic programming equation. This is so because of the "curse of dimensionality." The need, therefore, naturally arises for suboptimal algorithms. Those suboptimal algorithms should, however, share desirable qualitative features with the optimal controls. The study of simple examples of discrete-time linear systems with quadratic cost and multiplicative noise indicates two consequences of parameter uncertainty on the optimal control law. On the one hand, the presence of uncertainty in the parameter has a stimulating action on the control because a control exercised at a given time can improve the accuracy of future parameter estimates. This effect has been called loosely the probing effect of the control. On the other hand, the presence of uncertainty which cannot be reduced by the control has an inhibitory, loosely called the caution, effect of the control; the larger those irreducible uncertainties, the more attenuated the control should be. None of these consequences of uncertainties, the so-called dual effect, are captured by the naive "certainty equivalent" (CE) control law, which is obtained by setting all random parameters to their a priori mean values and treating the system as deterministic.

In the more general cases, wide-sense dual adaptive algorithms have been suggested ([1], [2], [3]). The crux of those adaptive algorithms is to approximate the cost-to-go in the dynamic programming equation by expanding it about a nominal trajectory to second-order terms in perturbations resulting from random disturbances. The resulting cost, called the dual cost, is minimized to yield the suboptimal control at the corresponding time stage. It has been observed by simulations [2], [5] that the algorithms displayed the desirable caution and probing features. Moreover, it has been claimed [5] that the dual cost could be decomposed in a sum of terms which account respectively for the caution effect, the probing effect and the deterministic part of the cost.

In general, however, it is impossible to compare such dual control laws with the optimal one, which is unknown, in the case of constant but unknown parameters. We consider here a different special case of a scalar, discrete-time linear system with white multiplicative gaussian noise and perfectly observed state. The optimal control law of such systems, for a quadratic performance index, is known [4], [7]. We show that, in that special case, it is possible to explicitly derive the dual cost and the dual control in closed form, when the length of the planning horizon goes to infinity. The dual cost on an infinite horizon is always finite, provided a simple controllability and positive definiteness assumption holds. This is a qualitative difference with respect to the optimal cost, which has been shown [4] to be infinite on an infinite horizon, unless some inequality is satisfied by the covariances; that property has been referred to as the <u>uncertainty threshold principle.</u> Thus, the dual control fails to exhibit that property.

Some valuable insight can be obtained, since we show that the asymptotic (i.e., infinite horizon) dual control law is in fact equivalent to a first-order expansion of the optimal control law for systems with white parameters as a function of the parameter covariances, about the nominal value of null parameter covariances, which corresponds to a deterministic problem. Since the certainty-equivalent (CE) control is simply a zero-order approximation, the dual control is shown to be intermediate (optimal to linear terms) between the CE and the optimal control. It is also understandable why the uncertainty threshold principle cannot be captured by the dual control, because it is an effect which is essentially nonlinear (quadratic and higher order terms) in the covariances. The accuracy of the dual control law for small parameter covariances is quite surprising, as no learning can take place in this problem, due to the white-noise parameter assumption. In other words, if the system parameters have small standard deviations about their mean values, we demonstrate by means of a scalar example that the dual control is (to first order linear terms in the parameter standard deviations) identical to the white-parameter optimal control law, which involves no learning. One can argue both ways whether this is "good news or bad news". The "good news" is that if the system parameters are not very random, then the inherent "robustness" properties of feedback, modulated correctly for parameter uncertainty, require no detailed "learning" of the parameters, provided that certain "caution" is exercised (this is not what the certainty-equivalence principle states). The "bad news" is that the dual control algorithm does not seem to capture the required "caution" effects when the system parmeters are very uncertain and very weakly correlated in time.

By the above comments we do not mean to imply any criticism of the dual adaptive control algorithm. It represents an excellent contribution to the state of the art in the field of stochastic adaptive control, and (again loosely speaking) it represents an intermediate approach to the control of systems with random parameters, somewhere between the case of perfect parameter knowledge assumptions (the certainty-equivalence case) and the (unrealistic) case that no learning of the system parameters is possible (the white multiplicative parameter case). What the authors attempt to do in this paper, by means of the simplest possible scalar example, is to understand some of the theoretical properties of the dual control algorithm. Thus, the reader should expect only a relatively minor theoretical contribution; by no means we imply any superiority of any stochastic adaptive control scheme that is useful for practical designs. The entire field of adaptive control has not yet matured to the point that can provide the engineering designer with useful instructions on how to realize an adaptive control system.

Another contribution of this paper is to examine the structure of the stochastic cost to go. In the dual control method the cost is split into three parts, the deterministic cost, the caution cost, and the probing cost. One would suspect that the probing part of the cost would correspond to the active learning of the unknown parameters, and that it would be zero in this example with multiplicative white parameters. However, the splitting of the dual cost between a caution and a probing term fails to have an appealing meaning. Both terms combine to yield a sum of positive weightings of the one-step predictions of the state covariances. Thus, no distinction can be made between uncertainties that can or cannot be influenced by the control.

In Section 2, the control problem is introduced. Section 3 presents its optimal solution on a finite horizon and discusses its existence on an infinite horizon, which is governed by the uncertainty threshold principle. In Section 4, the dual adaptive control algorithm is applied to the problem of concern. A closed form expression for the dual cost is derived and it is proved that it remains finite on an infinite horizon, under mild assumptions. In section 5, the comparison between the optimal, dual and certainty-equivalent law is performed for the infinite-horizon case. In section 6, the decomposition of the dual cost is examined. Section 7 contains the conclusion.

2. A Scalar, Multiplicative White-Noise Control Problem

The simple discrete-time stochastic control problem which will be considered here is the following:

$$x(k+1) = a(k)x(k) + b(k)u(k) \tag{2.1}$$

$$y(k) = x(k)$$

The state $x(k)$ is scalar and perfectly observed, without observation noise. There is also no additive process noise. The time constant $a(k)$ and the control gain $b(k)$ are unknown parameters. They are independent from one stage to another; namely, they constitute a white noise sequence. In addition, they are assumed to be gaussian, with means $\bar{a}$, $\bar{b}$ and covariance matrix

$$\underline{\Sigma} = \begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ab} & \Sigma_{bb} \end{bmatrix} \tag{2.2}$$

Therefore,

$$\begin{aligned} E\{[a(k)-\bar{a}]\,[b(j)-\bar{b}]\} &= \Sigma_{ab}\,\delta_{jk} \\ E\{[a(k)-\bar{a}]\,[a(j)-\bar{a}]\} &= \Sigma_{aa}\delta_{jk} \\ E\{[b(k)-\bar{b}]\,[b(j)-\bar{b}]\} &= \Sigma_{bb}\delta_{jk} \end{aligned} \tag{2.3}$$

where δ_{jk} is the Kronecker delta.

The initial state $x(0)$ is known and the cost function is quadratic:

$$J = E\left\{\sum_{k=0}^{N-1} [Q\,x^2(k) + Ru^2(k)] + Qx^2(N)\right\} \tag{2.4}$$

This stochastic control problem is one of the few which yield themselves to a closed-form analytical solution [4]. On the other hand, its structure is simple enough so that the dual cost can be expressed in closed-form, too, at least when the horizon length N is *infinite*. This makes a comparison possible between the optimal solution and the suboptimal solutions obtained from either the dual adaptive algorithm [2] or the certainty equivalent strategy.

3. Optimal Control: Finite- and Infinite-Horizon Cases

3.1 Finite horizon

Because of the Gaussian character of the random parameters, their probability distribution is entirely characterized by its first- and second-order moments. If $\bar{a}$ and $\bar{b}$ denote the expectations of a and b respectively, the optimal control law is found [4] to be the *following* linear feedback law:

$$u_{OPT}(j) = -\,G(j)x(j) \tag{3.1}$$

where

$$G(j) = \frac{K(j+1)\ (\Sigma_{ab} + \bar{a}\,\bar{b})}{R + K(j+1)(\Sigma_{bb} + \bar{b}^2)} \tag{3.2}$$

and the scalars K(j + 1) are given by a backward recursion of Riccati type:

$$K(j) = Q + K(j+1)(\Sigma_{aa} + \bar{a}^2) - \frac{K^2(j+1)(\Sigma_{ab} + \bar{a}\,\bar{b})^2}{R + K(j+1)(\Sigma_{bb} + \bar{b}^2)} \tag{3.3}$$

$$K(N) = 0$$

The optimal cost is equal to:

$$J^* = K(0)x^2(0) \tag{3.4}$$

Each K(j) can be viewed as a function of Σ_{aa}, Σ_{ab}, Σ_{bb}. The deterministic problem, where a and b are known parameters, corresponds to the values $\Sigma_{aa} = \Sigma_{ab} = \Sigma_{bb} = 0$. It can be verified indeed that, upon setting the covariances to zero in (3.3) and (3.2), the solution of the deterministic linear-quadratic problem is obtained. The certainty-equivalent (CE) control strategy consists of replacing the unknown parameters by their current estimates and then solving the corresponding deterministic control problem. In the present problem, because of the white-noise property, the best current estimates of a and b are their a priori means $\bar{a}$ and $\bar{b}$, since no learning is possible. Therefore, the certainty-equivalent strategy amounts to setting $\Sigma_{aa} = \Sigma_{ab} = \Sigma_{bb} = 0$ in (3.1), (3.2), (3.3). It is a zero-order approximation of K(j) as a function of $\underline{\Sigma}$ about $\underline{\Sigma} = 0$. Intuitively, the certainty-equivalent law will become increasingly poorer as the covariances depart further from zero, i.e., as the problem becomes more stochastic. This is evidently always so, but it will be seen below that the same remark applies to the dual adaptive control law as well.

3.2 Infinite horizon; the Uncertainty Threshold Principle

When the time horizon N goes to infinity, it can be shown [4] that the optimal cost J* need not remain bounded. In fact, a necessary and sufficient condition for the cost J* to go to a finite limit when $N \to \infty$ is that the following inequality between the covariances should hold:

$$\Sigma_{aa} + \bar{a}^2 - \frac{(\Sigma_{ab} + \bar{a}\,\bar{b})^2}{\Sigma_{bb} + \bar{b}^2} < 1 \tag{3.5}$$

The left-hand side of (3.5) has been called the uncertainty threshold, and the property, the uncertainty threshold principle. This is an essentially nonlinear result, which states that, if the covariance matrix $\underline{\Sigma}$ lies outside of a certain region, the asymptotic infinite horizon problem is ill-posed. This is in sharp contrast with the deterministic problem where, under mild controllability and positive-definiteness assumptions, the optimal cost reaches a finite limit as $N \to \infty$. In the present problem, those assumptions are:

$$\bar{b} \neq 0; \quad Q > 0, \quad R > 0. \tag{3.6}$$

It is seen that, for $\Sigma_{aa} = \Sigma_{ab} = \Sigma_{bb} = 0$, the left-hand side of (3.5) is equal to zero, so that (3.5) is satisfied.

When the inequality (3.5) is satisfied, the limit of $K(j)$ for $N \to \infty$ is the solution of the algebraic equation corresponding to (3.3), namely:

$$K = Q + K(\Sigma_{aa} + \bar{a}^2) - \frac{K^2(\Sigma_{ab} + \bar{a}\,\bar{b})^2}{R + K(\Sigma_{bb} + \bar{b}^2)} \tag{3.7}$$

Inequality (3.5) is also necessary and sufficient in order for the algebraic equation (3.7) to have a unique positive solution [4]. This solution will be denoted $\bar{K}$. It is in fact a function of the covariance matrix $\underline{\Sigma}$:

$$\bar{K} = \bar{K}(\underline{\Sigma}) .$$

An alternative way of stating the uncertainty threshold principle is therefore as follows.

The nonlinear function $\bar{K}(\underline{\Sigma})$ is defined on the region of the space of $\underline{\Sigma}$ described by (3.5), and it approaches infinitely as $\underline{\Sigma}$ goes to the boundary of that region. Note that the asymptotic value of the cost in the CE strategy is obtained from the value of $\bar{K}(\underline{\Sigma})$ at $\underline{\Sigma} = 0$, just as in the finite-horizon case.

4. Dual Adaptive Control

4.1 Expression for the Dual Cost

In this section, we now apply to the stochastic control problem introduced in section 2 the wide-sense dual adoptive control algorithm of Tse and Bar-Shalom [1], [2]. This algorithm consists of approximating the cost-to-go from step k + 1 on in the dynamic programming equation; the sum of the cost at stage k and of this approximated cost-to-go, called the dual cost, is minimized with respect to the control u(k) to yield the dual control at step k. The approximation of the cost-to-go is carried out in two steps. In the first step, the enlarged state $\underline{z}(k + 1)$ consisting of the initial state x(k) and the random parameter is estimated at time (k+1) from the information available at time k, and the optimal cost corresponding to a deterministic dynamical system is calculated. This deterministic dynamical system is obtained from the stochastic one by setting the random parameters to their expectations.

This step is essentially the application of the CE control, say $u_0(j)$, from time (k + 1) on. It results in a nominal trajectory $\underline{z}_0(j)$ $(j \geq k+1)$, and a nominal estimate $J_0(k + 1)$ of the cost-to-go.

In a second step, random disturbances $\xi(j)$ (for $j \geq k + 1$) are introduced; they cause perturbations $\delta\underline{z}(j)$ of the nominal states. The new trajectory is described by

$$\underline{z}(j) = \underline{z}_0(j) + \delta\underline{z}(j) \; ; \qquad (j \geq k+1) \; .$$

Perturbation controls $\delta u(j)$ are exercised so as to minimize the expected increment in the cost, $\Delta J_o(k+1)$. In order to solve that minimization problem, the state perturbation $\delta\underline{z}(j+1)$ is expanded to second-order terms in $\delta\underline{z}(j)$ and $\delta u(j)$, using the dynamical eqution about the nominal trajectory $\underline{z}_o(j)$ and control $u_o(j)$. The cost function is also expanded to second order about the nominal trajectory.

This permits the evaluation of $\Delta J_o^*(k+1)$, the minimum of $\Delta J_o(k+1)$. The wide-sense dual adaptive control a time k, $u_d(k)$, is then obtained by minimizing over the input $u(k)$ the dual cost $J_d[u(k)]$, namely the sum of the one-step cost at stage k and the approximation of the cost-to-go from stage $(k+1)$ on:

$$J_d[u(k)] \triangleq E\{Qx^2(k) + Ru^2(k) + J_o(k+1) + \Delta J_o^*(k+1) | Y_k\} \tag{4.1}$$

where Y_k denotes the information available at stage k. In the present problem, Y_k can be described by the sequences $x(i)$ $(i=0,1,\ldots,k-1)$ and $u(i)$ $(i=0,1,\ldots,k-1)$.

In the problem introduced in section 2 , the enlarged state is defined by

$$\underline{z}^T(k) = [x(k), a(k), b(k)] \tag{4.2}$$

Step 1 of the dual adaptive algorithm sets the initial state at time $(k+1)$ to the estimated value, given the information Y_k:

$$x_o(k+1) = \hat{x}(k+1|k) = \overline{a}x(k) + \overline{b}u(k) \tag{4.3}$$

The deterministic version of the dynamical equation (2.1) is

$$x_o(j+1) = \bar{a}x_o(j) + \bar{b}u_o(j) \qquad (j = k+1,\ldots,N-1) \tag{4.4}$$

The nominal control sequence is the optimal control sequence of the associated deterministic problem from time (k + 1) on:

$$u_o(j) = -G_o(j)x_o(j) = -\frac{\bar{K}_o(j+1)\bar{a}\,\bar{b}}{R + \bar{K}_o(j+1)\bar{b}^2}\,x_o(j) \tag{4.5}$$

and $\bar{K}_o(j)$ is given recursively by the Riccati difference equation:

$$\bar{K}_o(j) = Q + \bar{K}_o(j+1)\bar{a}^2 - \frac{\bar{K}_o^2(j+1)\bar{a}^2\bar{b}^2}{R + \bar{K}_o(j+1)\bar{b}^2} \tag{4.6}$$

$$\bar{K}_o(N) = 0$$

Equation (4.6) is in fact the special version of Eq. (3.3) corresponding to $\Sigma_{aa} = \Sigma_{ab} = \Sigma_{bb} = 0$. The initial estimate of the cost-to-go is given by:

$$J_o(k+1) = (1/2)\bar{K}_o(k+1)x_o^2(k+1) = (1/2)\bar{K}_o(k+1)[\bar{a}x(k) + \bar{b}u(k)]^2 \tag{4.7}$$

In step 2 of the dual adaptive algorithm, the covariances of the enlarged state appear, in the calculation of the cost perturbation $\Delta J_o^*(k+1)$. The updated covariance matrix of the perturbation $\delta\underline{z}(j)$ of the enlarged state $\underline{z}(j)$ given the current information, along the nominal trajectory, is:

$$\underline{\Sigma}_o(j|j) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \Sigma_{aa} & \Sigma_{ab} \\ 0 & \Sigma_{ab} & \Sigma_{bb} \end{bmatrix} \tag{4.8}$$

This results from the fact that the state x(k) is exactly observed and from the white noise assumption on a(k) and b(k). The one-step predicted covariance of the perturbation of the enlarged state, along the nominal trajectory, is

$$\underline{\Sigma}_o(j+1|j) = \begin{bmatrix} \Sigma^o_{xx}(j+1|j) & 0 & 0 \\ 0 & \Sigma_{aa} & \Sigma_{ab} \\ 0 & \Sigma_{ab} & \Sigma_{bb} \end{bmatrix} \tag{4.9}$$

where

$$\Sigma^o_{xx}(j+1|j) = \Sigma_{aa}x_o^2(j) + 2\Sigma_{ab}x_o(j)u_o(j) + \Sigma_{bb}u_o^2(j) \tag{4.10}$$

and $x_o(j)$, $u_o(j)$ denote the nominal trajectory and control, as obtained in step 1.

Equation (4.9) also results from the white-noise assumption and the perfect observation of the state. From the expression for the dual cost as given in Tse et al. ([3], Eq. (3-12)) it follows that

$$\begin{aligned} J_d[u(k)] &= (1/2)Ru^2(k) + (1/2)\overline{K}_o(k+1)\hat{x}(k+1|k)^2 + p_o(k+1)\hat{x}(k+1) \\ &+ (1/2)\mathrm{tr}\left\{\sum_{j=k+1}^{N} \underline{W}(j)\underline{\Sigma}_o(j|j) + [\underline{\Sigma}(k+1|k) - \underline{\Sigma}_o(k+1|k+1)]\underline{K}_o(k+1)\right. \\ &+ \left.\sum_{j=k+1}^{N-1} [\underline{\Sigma}_o(j+1|j) - \underline{\Sigma}_o(j+1|j+1)]\underline{K}_o(j+1)\right\} \end{aligned} \tag{4.11}$$

In Eq. (4.11), $\underline{K}_o(j)$ is a matrix which has the dimension of the enlarged state. Denoting the random parameters by the vector

$$\underline{\theta}(k) = [a(k), b(k)] , \tag{4.12}$$

the matrix $\underline{K}_o(j)$ can be partitioned as

$$\underline{K}_o(j) = \begin{bmatrix} K_o^{xx}(j) & \underline{K}_o^{x\theta}(j) \\ \underline{K}_o^{\theta x}(j) & \underline{K}_o^{\theta\theta}(j) \end{bmatrix} \tag{4.13}$$

It turns out [3] that

$$K_o^{xx}(j) = \bar{K}_o(j) \tag{4.14}$$

where $\bar{K}_o(j)$ is the solution of the Riccati difference equation (4.6).

The matrices $\underline{K}_o^{x\theta}(j)$ and $\underline{K}_o^{\theta\theta}(j)$ can be obtained from recursions [3] once the sequence $K_o^{xx}(j)$ is known. The vector $\underline{p}_o(k+1)$ is zero in our example because we deal with a regulator, not a tracking problem. Also, $\underline{\Sigma}(k+1|k)$, the one-step predicted covariance of the enlarged state at stage k, is given by (4.9) with j=k, since

$$\hat{x}(k+1|k) = \bar{a}\, x(k) + \bar{b}u(k) \tag{4.15}$$

In eq. (4.11), the matrix $\underline{W}(j)$ has the following structure (see [3], Eq. (3.17)):

$$\underline{W}(j) = \begin{bmatrix} Q & V_1 & V_2 \\ V_1 & 0 & 0 \\ V_2 & 0 & 0 \end{bmatrix} \tag{4.16}$$

The exact definition of V_1, V_2 is unimportant in this example, because

$$\mathrm{tr}[\underline{W}(j)\underline{\Sigma}_o(j|j)] = \mathrm{tr}\left\{\begin{bmatrix} Q & V_1 & V_2 \\ V_1 & 0 & 0 \\ V_2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & \Sigma_{aa} & \Sigma_{ab} \\ 0 & \Sigma_{ab} & \Sigma_{bb} \end{bmatrix}\right\} = 0 \tag{4.17}$$

On the other hand,

$$\mathrm{tr}\left\{\underline{K}_o(j+1)\left[\underline{\Sigma}_o(j+1|j) - \underline{\Sigma}_o(j|j)\right]\right\} = \mathrm{tr}\begin{bmatrix} \Sigma_{xx}^o(j+1|j)\;\; K_o^{xx}(j+1) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \Sigma_{xx}^o(j+1|j)K_o^{xx}(j+1) = \Sigma_{xx}^o(j+1|j)\bar{K}_o(j+1) \tag{4.18}$$

From Eqs. (4.11), (4.14), (4.15), (4.17), (4.18), and the remarks just made, it follows that the dual cost is given as follows in our problem.

$$J_d[u(k)] = (1/2)Ru^2(k) + (1/2)[\bar{a}x(k) + \bar{b}u(k)]^2\bar{K}_o(k+1) + (1/2)\bar{K}_o(k+1)\Sigma_{xx}(k+1|k) + (1/2)\sum_{j=k+1}^{N-1}\bar{K}_o(j+1)\Sigma^o_{xx}(j+1|j) \tag{4.19}$$

4.2 Infinite-Horizon Case

It will now be shown that, under the same assumptions which guarantee the finiteness of the certainty-equivalent cost over an infinite horizon, the dual cost too remains bounded. Therefore, there is a qualitative difference between the dual adaptive control and the optimal control in the infinite-horizon case: the former does not obey the uncertainty threshold principle, which governs the latter.

Controllability of the deterministic dynamical system (4.4) is equivalent to the property that $\bar{b} \neq 0$. Under the assumptions (3.6) ($\bar{b} \neq 0$, $Q > 0$, $R > 0$), it is well known [8] that the solution $\bar{K}_o(j)$ of Riccati recursion (4.6) reaches a finite positive limit $\bar{K}_o$ as $N \to \infty$. Hence, if $u_o(j)$, $x_o(j)$ are respectively successive controls and states of the certainty-equivalent strategy,

$$\lim_{N\to\infty}\sum_{j=k+1}^{N-1}[Qx_o^2(j) + Ru_o^2(j)] = (1/2)\bar{K}_o x_o^2(k+1) < +\infty \tag{4.20}$$

According to Eq. (4.19), we must prove that

$\sum_{j=k+1}^{N-1}\bar{K}_o(j+1)\Sigma^o_{xx}(j+1|j)$ remains bounded as $N \to \infty$. From Eqs. (4.5) and (4.10),

$$\sum_{j=k+1}^{N-1} \bar{K}_o(j+1)\Sigma^o_{xx}(j+1|j) = \sum_{j=k+1}^{N-1} \bar{K}_o(j+1)\,[\Sigma_{aa} - 2\,\Sigma_{ab}G_o(j) + \Sigma_{bb}G_o^2(j)]\,x_o^2(j) \tag{4.21}$$

In fact, both $\bar{K}_o(j+1)$ and $G_o(j)$ depend on N: let us emphasize that dependence by the notation $\bar{K}_o(j+1;\,N)$, $G_o(j;\,N)$. Clearly,

$$\bar{K}_o(j+1;\,N) \le \bar{K}_0(j;\,N) \tag{4.22}$$

since the left-hand side defines the minimal cost on a shorter horizon. From Eq. (4.5),

$$\frac{\partial G_o(j)}{\partial \bar{K}_o(j+1)} = \frac{\bar{a}\,\bar{b}\,R}{[R+\bar{K}_o(j+1)\bar{b}^2\,]^2}$$

Accordingly

$$\frac{\partial G_o^2(j)}{\partial \bar{K}_o(j+1)} = 2G_o(j)\,\frac{\partial G_o(j)}{\partial \bar{K}_o(j+1)} = \frac{R(\bar{a}\,\bar{b})^2\,K_o(j+1)}{[R + \bar{K}_o(j+1)\bar{b}^2\,]^3} \ge 0$$

whence it follows that, also,

$$G_o^2(j+1;N) \le G_o^2(j\,;\,N)$$

or

$$|G_o(j+1;N)| \le |G_o(j;N)| \tag{4.23}$$

From (4.22) and (4.23),

$$\bar{K}_o(j+1;N) \le \bar{K}_o$$

and

$$|G_o(j+1;N)| \le |G_o|$$

where G_o is obtained from $\bar{K}_o$ by the same function which yields $G_o(j)$ from $\bar{K}_o(j)$. (Eq. (4.5)). On the other hand, $\underline{\Sigma}$, as a covariance matrix, is symmetric and positive semi-definite. Therefore,

$$[\Sigma_{aa}-2\Sigma_{ab}G_o(j) + \Sigma_{bb}G_o^2(j)] \leq \sigma[G_o^2(j) + 1] \tag{4.24}$$

where σ is the largest eigenvalue of $\underline{\Sigma}$. As a result,

$$\sum_{j=k+1}^{N-1} \bar{K}_o(j+1)\, \Sigma^o_{xx}(j+1|j) \leq \bar{K}_o\sigma[G_o^2 + 1] \sum_{j=k+1}^{N-1} x_o^2(j) \tag{4.25}$$

However,

$$Q \lim_{N\to\infty} \left[\sum_{j=k+1}^{N-1} x_o^2(j)\right] \leq \lim_{N\to\infty} \sum_{j=k+1}^{N-1} [Qx_o^2(j) + Ru_o^2(j)] = 1/2\ \bar{K}_o x_o^2(k+1) < +\infty$$

Since $Q > 0$, it follows that

$$\lim_{N\to\infty} \sum_{j=k+1}^{N-1} x_o^2(j) < +\infty$$

and therefore, the left-hand side of (4.25) remains finite as $N \to \infty$.

A consequence of this observation is that there will be an important discrepancy between the trajectory resulting from the application of the dual control, and the optimal trajectory, for the range of covariances which do not obey the inequality (3.5) of the uncertainty threshold principle. This qualitative difference is confirmed by a quantitative comparison in the next section.

5. Comparison between optimal and dual control in the infinite-horizon Case

From the expression (4.19) for the dual cost and the knowledge that it remains bounded on an infinite horizon (section 4.2), it is possible to obtain a closed-form expression for the limit of the dual cost when N goes to infinity, in terms of the various problem data and the limit $\bar{K}_o$ of the solution to the Riccati recursion (4.6). This in turn provides a closed-form expression for the dual adaptive control, which can therefore be compared with the optimal control as given by Eqs. (3.1), (3.2), (3.3). Let

$$a(j;N) \stackrel{\Delta}{=} \bar{K}_o(j+1;N)\, \Sigma^o_{xx}(j+1|j) \tag{5.1}$$

We are interested in evaluating

$$L = \lim_{N\to\infty} \sum_{j=k+1}^{N} a(j,N) \tag{5.2}$$

To that end, we use Eq. (4.22), and the stability of the closed-loop dynamical system of the certainty-equivalent strategy. Namely,

$$x_o(j+1) = \Delta_o(j)x_o(j) \tag{5.3}$$

where

$$\Delta_o(j) = \bar{a} - \bar{b}\, G_o(j) = \frac{\bar{a}\, R}{R + \bar{K}_o(j+1)\bar{b}^2} \tag{5.4}$$

It is known [8] that, under the assumptions (3.6), the asymptotic closed-loop system is strictly stable:

$$|\Delta| = \lim_{N\to\infty}|\Delta_o(j)| = \frac{|\bar{a}|R}{R+\bar{K}_o\bar{b}^2} < 1 \tag{5.5}$$

From eqs. (4.21) and (5.3),

$$a(j,N) = \bar{K}_o(j+1)[\Sigma_{aa} - 2\Sigma_{ab}G_o(j) + \Sigma_{bb}G_o^2(j)] \prod_{i=k+1}^{j-1} \Delta^2(i)x^2(k+1) \tag{5.6}$$

where, in fact, $\bar{K}_o(j+1) = \bar{K}_o(j+1;N)$, $G_o(j) = G_o(j;N)$, $\Delta(i) = \Delta(i;N)$.

However,

$$L = \lim_{N\to\infty} \sum_{m=k+1}^{N} a(j,N) = \lim_{N\to\infty} \left[\sum_{j=k+1}^{m} a(j,N) + \sum_{j=m+1}^{N} a(j,N) \right]$$

for any $m \in \{k+1,\ldots,N-1\}$. Therefore, also,

$$L = \lim_{m\to\infty} \lim_{N\to\infty} \left[\sum_{j=k+1}^{m} a(j,N) + \sum_{j=m+1}^{N} a(j,N) \right]$$

But, it has been shown in section 4.2 that

$$\lim_{N\to\infty} \sum_{j=k}^{N} a(j,N) < +\infty \qquad \text{for all } k$$

Accordingly,

$$\lim_{m\to\infty} \left[\lim_{N\to\infty} \sum_{j=m+1}^{N} a(j,N) \right] = 0 \quad .$$

and

$$L = \lim_{m\to\infty} \lim_{N\to\infty} \sum_{j=k+1}^{m} a(j,N) = \lim_{m\to\infty} \left[\sum_{j=k+1}^{m} \lim_{N\to\infty} a(j,N) \right] \tag{5.7}$$

From (5.6), using the convergence of $K(j+1;N)$, $G(j;N)$ and $D(j;N)$, it is concluded that $a(j,N)$ goes to a limit as $N\to\infty$, and

$$\lim_{N\to\infty} a(j,N) = \bar{K}_o[\Sigma_{aa} - 2\,\Sigma_{ab}G_o + \Sigma_{bb}G_o^2]\Delta^{2(j-1-k)}x^2(k+1) \tag{5.8}$$

where

$$G_o = \frac{\bar{K}_o\,\bar{a}\,\bar{b}}{R + \bar{K}_o\,\bar{b}^2}$$

Accordingly,

$$L = \bar{K}_o[\Sigma_{aa} - 2\,\Sigma_{ab}\,G_o + \Sigma_{bb}G_o^2]\; x_0^2(k+1) \lim_{m\to\infty} \sum_{j=k+1}^{m} \Delta^{2(j-1-k)}$$

$$= \frac{\bar{K}_o}{1-\Delta^2}\,[\Sigma_{aa} - 2\,\Sigma_{ab}G_o + \Sigma_{bb}\,G_o^2\,]\,x_0^2(k+1) \tag{5.9}$$

where the second equality results from $|\Delta| < 1$.

In summary, taking into account Eqs. (4.3), (4.19), and (5.9), the asymptotic value of the dual cost is arrived at:

$$\begin{aligned} \lim_{N\to\infty} J_d[u(k)] = {} & (1/2)Ru^2(k) + (1/2)\bar{K}_o[\bar{a}\,x(k) + \bar{b}u(k)]^2 \\ & + (1/2)\bar{K}_o[\Sigma_{aa}x^2(k) + 2\Sigma_{ab}x(k)u(k) + \Sigma_{bb}u(k)] \\ & + (1/2)\,\frac{\bar{K}_o}{1-\Delta^2}\,[\Sigma_{aa} - 2G_o\Sigma_{ab} + G_o^2\Sigma_{bb}][\bar{a}x(k) + \bar{b}u(k)]^2 \end{aligned} \tag{5.10}$$

The minimization of the asymptotic dual cost (5.10) with respect to u(k) yields the stationary dual adaptive control law, $u_d[x(k)]$.

$$u_d[x(k)] = -\frac{\bar{K}_o\Sigma_{ab} + [\bar{K}_o + \frac{\bar{K}_o}{1-\Delta^2}(\Sigma_{aa} - 2G_o\Sigma_{ab} + G_o^2\Sigma_{bb})]\bar{a}\,\bar{b}}{R+\bar{b}^2\left[\bar{K}_o + \frac{\bar{K}_o}{1-\Delta^2}(\Sigma_{aa} - 2G_o\Sigma_{ab} + G_o^2\,\Sigma_{bb})\right] + \bar{K}_o\Sigma_{bb}}\,x(k) \tag{5.11}$$

Comparison of Eq. (5.11) with the asymptotic version of the optimal control law (3.2) evidences a similar structure. However, the limit of K(j+1) as $N\to\infty$ which occurs in (3.2) - if it exists - is the positive solution of Eq. (3.7) - if it exists; that is, $\bar{K}(\underline{\Sigma})$. Recall that that limit exists if and only if $\underline{\Sigma}$ lies within a region defined by Eq. (3.5). In contrast, the parameter $\bar{K}_o$ which occurs in (5.11) is always defined, finite and positive, and

$$\bar{K}_o = \bar{K}(\underline{\Sigma})\big|_{\underline{\Sigma}=o} \tag{5.12}$$

From Eq. (3.7), the gradient of $\bar{K}(\underline{\Sigma})$ with respect to $\underline{\Sigma}$, evaluated at $\underline{\Sigma}=0$, can be found (see appendix), and the resulting first-order expression of $\bar{K}(\underline{\Sigma})$ about $\underline{\Sigma}=0$ is accordingly found:

$$\bar{K}(\underline{\Sigma}) = \bar{K}_o + \frac{\partial K^T}{\partial \underline{\Sigma}} (\underline{\Sigma}) \Big|_o \underline{\Sigma} + o(\underline{\Sigma}) \tag{5.13}$$

and

$$\frac{\partial K^T}{\partial \underline{\Sigma}} (\underline{\Sigma}) \Big|_o \underline{\Sigma} = \frac{\bar{K}_o}{1-\Delta^2} (\Sigma_{aa} - 2G_o \Sigma_{ab} + G_o^2 \Sigma_{bb}) \tag{5.14}$$

Hence, the expression between brackets in (5.11) is recognized as the first-order expansion of $K(\Sigma)$ in $\underline{\Sigma}$ about $\underline{\Sigma} = 0$. Note that, from (5.14), $[\bar{K}_o + \frac{\partial K^T}{\partial \underline{\Sigma}} (\underline{\Sigma}) \big|_o \underline{\Sigma}]$ is positive, regardless of the value of the covariance matrix $\underline{\Sigma}$. This follows from (5.14), the fact that $\bar{K}_o > 0$, $1-\Delta^2 > 0$ and the positive semidefiniteness of $\underline{\Sigma}$. The stationary optimal control law (from Eqs. (3.1), (3.2)) exists in the neighborhood of $\underline{\Sigma}=0$ (because $\underline{\Sigma}=0$ satisfied Eq. (3.5) and by continuity) and can also be expanded to first order in $\underline{\Sigma}$:

$$u_{OPT}(\underline{\Sigma}) = u_{OPT}(0) + \left(\frac{\partial u_{OPT}}{\partial \underline{\Sigma}} \Bigg|_o + \frac{\partial u_{OPT}}{\partial \bar{K}} \frac{\partial K}{\partial \underline{\Sigma}} \Bigg|_o \right)^T \underline{\Sigma} + o(\underline{\Sigma}) \tag{5.15}$$

where both the direct dependence of u_{OPT} on $\underline{\Sigma}$ and the indirect dependence through $\bar{K}(\underline{\Sigma})$ have been taken into account. It follows from (5.15) and (5.13), (5.14) that

$$u_d(\underline{\Sigma}) - u_{OPT}(\underline{\Sigma}) = o(\underline{\Sigma})$$

or

$$\lim_{\underline{\Sigma}\to 0} \frac{|u_d(\underline{\Sigma}) - u_{OPT}(\underline{\Sigma})|}{||\underline{\Sigma}||} = 0 \tag{5.16}$$

where $||\underline{\Sigma}||$ is the euclidean norm, for instance. (See the appendix for a proof).

On the other hand, the approximation is no better than the first order. Indeed (see appendix), the second derivatives of u_{OPT} with respect to $\underline{\Sigma}$ involve the second derivatives of $K(\underline{\Sigma})$, evaluated at $\Sigma=0$, which are not present in the dual u_d. Therefore, the stationary dual control (on an infinite time-horizon) is the first-order approximation of the optimal control, as a function of the covariance matrix $\underline{\Sigma}$, about the numerical value $\underline{\Sigma}=0$ which corresponds to a deterministic problem. It has already been pointed out (section 3.1) that the certainty-equivalent control is a zeroth-order approximation, in the sense that

$$u_{CE}(\underline{\Sigma}) = u_{OPT}(\underline{\Sigma})\Big|_{\underline{\Sigma}=0}$$

This is apparent from Eq. (4.5). Thus, the result of this section shows that, in our particular problem and for an infinite horizon, the dual control performs better than the CE control, but less well than the optimal one. The accuracy of the dual control can be quite high for small covariances, which is somewhat surprising in view of the fact that the parameter cannot be learned, due to the white-noise property.

When the parameter covariances grow large, however, the discrepancy between the dual and the optimal control can become substantial. This is confirmed for instance by the consideration of limiting cases. Assume, for instance, that a and b are uncorrelated, with the variance of a being fixed. For $\Sigma_{bb}\to\infty$, $\overline{K}(\underline{\Sigma})$ goes to:

$$\overline{K}(\Sigma_{aa}) = \frac{Q}{1 - \overline{a}^2 - \Sigma_{aa}}$$

as is apparent from Eq. (3.7). The inequality (3.5) to be satisfied by the covariances is

$$\Sigma_{aa} + \bar{a}^2 < 1$$

The optimal control law (3.2) goes to zero when Σ_{bb} goes to infinity. This is an example of caution effect: the control is inhibited by uncertainties that it cannot affect. In constrast, the dual control law $u_d(k)$ goes to a finite limit

$$\lim_{\Sigma_{bb}\to\infty} u_d(k) = - \frac{\bar{a}\,\bar{b}}{\bar{b}^2 + \frac{(1-\Delta^2)}{G_o^2}} x(k)$$

Hence, one can say in that case that the dual law is not cautious enough. The fit can, however, sometimes be better, even at large values of the covariances. For instance, in another limiting case where a and b are still uncorrelated, but Σ_{bb} remains fixed and $\Sigma_{aa}\to\infty$, both laws have the same limit:

$$\lim_{\Sigma_{aa}\to\infty} u_{OPT}(k) = \lim u_d(k) = - \frac{\bar{a}}{\bar{b}} x(k) .$$

6. Decomposition of the Dual Cost

A decomposition of the dual cost for the general discrete stochastic control problem with quadratic cost, linear dynamical equations and linear evolution equation for the random parameters has been proposed in the literature [5]. This decomposition splits the dual cost into a deterministic term, a "caution" term and a "probing" term. The deterministic term $J_D(k)$ represents the value of the cost-to-go corresponding to the certainty-equivalent strategy, namely, it depends on the unknown coefficients only through their current estimated expectations.

The caution term $J_c(k)$ is supposed to reflect those uncertainties that the control at stage k cannot affect directly, although it can affect their weightings. Those include the one-step predicted covariance of the enlarged state at stage k, and the covariance of the noise of the enlarged state.

The probing term $J_p(k)$ contains those uncertainties which the control at stage k can influence; those include the future updated covariances. In our problem however, the updated covariance matrices of future states are all equal to the a priori covariance matrix of the paramters a, b, because of the white-noise property, so that they cannot be influenced by the control. The various components of the dual cost are as follows [5], [6]:

$$J_D(k) = (1/2)Ru^2(k) + (1/2)[\bar{a}x(k) + \bar{b}u(k)]^2\bar{K}_o(k+1) \tag{6.1}$$

$$J_c(k) = (1/2)\bar{K}_o(k+1)\,\Sigma^{xx}_{k+1|k} + (1/2)\sum_{j=k}^{N-1} \mathrm{tr}\,(\underline{K}^{\theta\theta}_{j+1}\Sigma^{o|\theta\theta}_{j|j}) \tag{6.2}$$

and $J_p(k)$ is given [5] as a function of $\underline{K}^{\theta\theta}_{j+1}$ and K^{xx}_{j+1}, for $j = k+1,\ldots,N-1$. The dual cost is the sum of the three terms:

$$J_d(k) = J_D(k) + J_c(k) + J_p(k)\ . \tag{6.3}$$

Using the recursions [3] satisfied by $\underline{K}^{\theta\theta}_{j+1}$, and Eq. (4.10) for $\Sigma^o_{xx}(j+1|j)$, it is possible to verify that (6.3) is consistent with (4.19). However, the caution and probing terms combine, in our example, to yield:

$$J_c(k) + J_p(k) = (1/2)\bar{K}_o(k+1)\Sigma_{xx}(k+1|k) + (1/2) \sum_{j=k+1}^{N-1} \bar{K}_o(j+1)\Sigma^o_{xx}(j+1|j) \tag{6.4}$$

In view of Eq. (4.10), it is clear that the control u(k) can affect both $\Sigma_{xx}(k+1|k)$ and $\Sigma^o_{xx}(j+1|j)$, for $j \geq k+1$, but it cannot affect the coefficients $\bar{K}_o(i)$ $(i=k+1,\ldots,N)$. Hence, the decomposition into (6.1) and (6.2) does not seem to have any intuitive appeal in the present situation.

Perhaps, another splitting of the cost would be more appropriate, where the nondeterministic part of the cost, J_d-J_D, would be expressed as the sum of one term which corresponds to the open-loop feedback strategy [7], and the difference.

In conclusion, it seems that, even though the dual algorithm is very near optimality for small covariances (Section 5), its action cannot be explained by the decomposition between probing and caution in the present scalar example.

7. Conclusion and Suggestions for Future Work

The motivation for this analysis has been the desire to gain more insight into the behavior of the wide-sense dual control algorithm [1], [2], whose available results so far arise from simulations. Those results are, of necessity, qualitative rather than quantitative because a comparison of the adaptive control with the optimal control is usually impossible since the latter is unknown. An attempt towards the quantization of some desirable adaptive features possessed by the dual control - probing and caution - was made recently [5], by splitting the dual cost into component terms which each are claimed to account for a particular effect.

The approach that we have taken here has been to concentrate on a special discrete stochastic control problem (quadratic cost, linear dynamics, multiplicative guassian white noise with perfectly observed state) where the optimal control is known. The special nature of the problem makes it possible to evaluate the dual control, too, in closed analytical form, at least for the infinite-hroizon case. This permits a thorough comparison with the optimal control, which reveals (1) that the dual control does not share a fundamental property of the optimal control, the uncertainty threshold principle; (2) that the dual control approximates the optimal control linearly in the covariances of the random parameters, for small values of the parameter covariances.

Since no learning can occur (because the parameters are white-noise), one would expect the probing term in the dual cost to vanish. This is, however, not the case. Instead, probing term and caution term combine to yield a positively weighted sum of the one-step predicted covariances of the future states. This observation makes one doubt the usefulness of the splitting between caution and probing terms in general, as we'l as their intuitive meaning.

Also, alternative decompositions of the dual cost should be investigated. Ideally, one term should correspond to the certainty-equivalent control law (this is accomplished by the deterministic term); another term, to the open-loop feedback law, and the remaining term would account for the learning characteristics of the algorithm.

8. References

1. E. Tse, Y. Bar-Shalom, L. Meier, III, "Wide-Sense Adaptive Dual Control for Nonlinear Stochastic Systems", IEEE Transactions on Automatic Control, Vol. AC-18, No. 2, April 1973.

2. E. Tse, Y. Bar-Shalom, "An Actively Adaptive Control for Linear Systems with Random Parameters via the Dual Control Approach", IEEE Transactions on Automatic Control, Vol. AC-18, No. 2, April 1973.

3. Y. Bar-Shalom, E. Tse, "Dual Effect, Certainty Equivalence, and Separation in Stochastic Control, IEEE Transactions on Automatic Control, Vol. AC-19, No. 5, October 1974.

4. M. Athans, R. Ku, S.B. Gershwin, "The Uncertainty Threshold Principle: Some Fundamental Limitations on Optimal Decision Making under Dynamic Uncertainty", IEEE Transactions on Automatic Control, Vol. AC-22, No. 3, June 1977.

5. Y. Bar-Shalom, K.D. Wall, "Effects of Uncertainties on the Adaptive Control of Macroeconomic Systems, Proceedings of IFAC Congress, Helsinki, 1978.

6. D. Kendrick, et al, "Control Theory and Econometric Models", to be published.

7. D. Bertsekas, "Dynamic Programming and Stochastic Control", Acad. Press, 1976.

8. P. Dorato, A.H. Levis, "Optimal Linear Regulators: the Discrete-Time Case", IEEE Transactions on Automatic Control, Vol. AC-16, December 1971.

9. Appendix

We shall establish equations (5.14) and (5.16).

1. Proof of Equation (5.14)

Equation (3.7) can be described abstractly as

$$F(K,\underline{\Sigma}) = 0 \tag{A.1}$$

Hence, if $K(\underline{\Sigma})$ denotes the positive solution, the following is an identity in $\underline{\Sigma}$:

$$F[K(\underline{\Sigma}),\underline{\Sigma}] = 0 \tag{A.2}$$

Upon differentiating (A.2) with respect to $\underline{\Sigma}$ about $\underline{\Sigma} = 0$, one obtains:

$$\left(\frac{\partial F}{\partial \underline{\Sigma}}\right)_o^T + \left(\frac{\partial F}{\partial K}\right)_o \left(\frac{\partial F}{\partial \underline{\Sigma}}\right)_o^T = 0 \tag{A.3}$$

whence

$$\left(\frac{\partial K}{\partial \underline{\Sigma}}\right)_o^T \underline{\Sigma} = - \frac{\left(\frac{\partial F}{\partial \underline{\Sigma}}\right)_o^T \underline{\Sigma}}{\left(\frac{\partial F}{\partial K}\right)_o} \qquad \text{(A.4)}$$

The numerator and the denominator in (A.4) are now calculated. From (3.7),

$$F(K,\underline{\Sigma}) = Q + K(\Sigma_{aa} + \bar{a}^2) - \frac{K^2(\Sigma_{ab} + \bar{a}\,\bar{b})^2}{R + K(\Sigma_{bb} + \bar{b}^2)} - K \qquad \text{(A.5)}$$

Denominator of (A.4)

$$\frac{\partial F}{\partial K}(K,\underline{\Sigma}) = \Sigma_{aa} + \bar{a}^2 - 1 - \frac{K(\Sigma_{ab} + \bar{a}\,\bar{b})^2[2R + 2K(\Sigma_{ab} + \bar{b}^2) - K(\Sigma_{bb} + \bar{b}^2)]}{[R + K(\;_{bb} + \bar{b}^2)]^2}$$

Therefore,

$$\left(\frac{\partial F}{\partial K}\right)_o = \bar{a}^2 - 1 - \frac{K\,\bar{a}^2\,\bar{b}^2(K\,\bar{b}^2 + 2R)}{(R + K\,\bar{b}^2)^2} = \bar{a}^2 - 1 - \frac{\bar{a}^2[(R + K\,\bar{b}^2)^2 - R^2]}{(R + K\,\bar{b}^2)^2}$$

$$= \frac{a^2R^2}{(R + K\,\bar{b}^2)^2} - 1 = \Delta^2 - 1 \qquad \text{(A.6)}$$

where the last equality results from (5.4), (5.5).

Numerator of (A.4)

$$\left(\frac{\partial F}{\partial \underline{\Sigma}}\right)_o^T \underline{\Sigma} = \left(\frac{\partial F}{\partial \Sigma_{aa}}\right)_o \Sigma_{aa} + \left(\frac{\partial F}{\partial \Sigma_{ab}}\right)_o \Sigma_{ab} + \left(\frac{\partial F}{\partial \Sigma_{bb}}\right)_o \Sigma_{bb} \qquad \text{(A.7)}$$

$$\frac{\partial F}{\partial \Sigma_{aa}} = \bar{K}$$

hence

$$\left(\frac{\partial F}{\partial \Sigma_{aa}}\right)_o = \bar{K} \tag{A.8}$$

$$\frac{\partial F}{\partial \Sigma_{bb}} = \frac{K^3(\Sigma_{ab} + \bar{a}\,\bar{b})^2}{[R + K(\Sigma_{bb} + \bar{b}^2)]^2}$$

$$\left(\frac{\partial F}{\partial \Sigma_{bb}}\right)_o = \frac{\bar{K}^3\,\bar{a}^2\,\bar{b}^2}{(R + \bar{K}\,\bar{b}^2)^2} \tag{A.9}$$

$$\frac{\partial F}{\partial \Sigma_{ab}} = \frac{-2\,K^2(\Sigma_{ab} + \bar{a}\,\bar{b})}{R + K(\Sigma_{bb} + \bar{b}^2)}$$

$$\left(\frac{\partial F}{\partial \Sigma_{ab}}\right)_o = \frac{-2\,\bar{K}^2\,\bar{a}\,\bar{b}}{R + \bar{K}\,\bar{b}^2} \tag{A.10}$$

Accordingly,

$$\left(\frac{\partial F}{\partial \underline{\Sigma}}\right)_o^T \underline{\Sigma} = \bar{K}[\Sigma_{aa} - \frac{2\,\bar{K}\,\bar{a}\,\bar{b}\,\Sigma_{ab}}{(R + \bar{K}\,\bar{b}^2)} + \frac{K^2\,\bar{a}^2\,\bar{b}^2}{(R + \bar{K}\,\bar{b}^2)^2}\,\Sigma_{bb}] \tag{A.11}$$

$$= K(\Sigma_{aa} - 2G\Sigma_{ab} + G^2\Sigma_{bb})$$

where Eq. (4.5) has been used. Thus, (5.14) results from (A.4) and (A.6), (A.11).

2. Proof of Equation (5.16)

The optimal control gain (3.1), (3.2) is a rational fraction in $K(\underline{\Sigma})$, more specifically a function of the type

$$u_{OPT}(\underline{\Sigma}) = \frac{A + BK(\underline{\Sigma}) + (\underline{c}^T\underline{\Sigma})K(\underline{\Sigma})}{A_1 + B_1K(\underline{\Sigma}) + (\underline{c}_1^T\underline{\Sigma})\;K(\underline{\Sigma})}\;x \tag{A.12}$$

On the other hand, the dual law is expressed by (5.11) as

$$u_d(\underline{\Sigma}) = \frac{A + B[K(0) + \underline{K}'(0)^T\underline{\Sigma}] + \underline{c}^T\underline{\Sigma}K(0)}{A_1 + B_1[K(0) + \underline{K}'(0)^T\underline{\Sigma}] + \underline{c}_1^T\underline{\Sigma}K(0)}\, x \tag{A.13}$$

where

$$\underline{K}'(0) \triangleq \left(\frac{\partial K}{\partial \underline{\Sigma}}\right)_o$$

with the same coefficients A, B, $\underline{c}$, A_1, B_1, $\underline{c}_1$ as in (A.12). Comparison of (A.12) and (A.13) shows that, in (A.13), both the numerator and the denominator of (A.12) have been replaced by their first expansion in $\underline{\Sigma}$ about 0. It follows that (A.12) and (A.13) have the same first-order expansion in $\underline{\Sigma}$ about $\underline{\Sigma} = 0$.

In effect,

$$u_{OPT}(0) = u_d(0) = \frac{A + B\,K(0)}{A_1 + B_1K(0)}\, x$$

and

$$\left(\frac{\partial u_{OPT}}{\partial \underline{\Sigma}}\right)_0 = \left(\frac{\partial u_d}{\partial \underline{\Sigma}}\right)_o = \frac{[A_1 + B_1K(0)][B\underline{K}'(0) + K(0)\underline{c}^T]}{[A_1 + B_1K(0)]^2}\, x - \frac{[A + BK(0)][B_1\underline{K}'(0) + K(0)\underline{c}_1^T]}{[A_1 + B_1K(0)]^2}\, x$$